# Strategies for Weather-Dependent Data Acquisition

E. C. Posner
TDA Planning Office
and
Electrical Engineering Department,
California Institute of Technology

*A strategy for data acquisition from a very distant spacecraft is presented, when the system performance can be severely degraded by the Earth's weather due to the high microwave frequency being used. Two cases are considered, one in which there is a certain minimum data rate to be maintained and one in which there isn't. The goal is to maximize expected data return, where we assume that there is always new data, or a backlog of old data, that can be sent if conditions are favorable. When there is no minimum rate to be maintained, the optimum strategy is the greedy strategy, which always transmits at that single rate which maximizes the expected data returned. If there is a minimum data rate that we strive to maintain even in adverse conditions, the optimum strategy transmits simultaneously at the minimum or base data rate and at a bonus data rate. We use a coding system designed for the bandwidth-constrained degraded broadcast channel. The optimum version of this system can, under realistic assumptions, save on the order of 5 dB over the conservative strategy of just transmitting at a single lower data rate.*

## I. Introduction

The Solar System Exploration Program uses X-band (8.5 GHz) to get a narrow antenna beam from spacecraft to Earth and so to maintain a high data rate in clear weather. And someday $K_a$-band (32 GHz) may be used. When reception is limited to receivers buried in the Earth's atmosphere, we must suffer weather effects which reduce the data rate from the value that we can maintain in clear dry weather. How can we best cope with this?

We assume a spacecraft so distant that the data rate cannot be changed to take account of the actual weather at a receiving station. At most, the spacecraft knows a statistical prediction of the weather that will be experienced at reception on Earth, and chooses a data rate or data rates based on this. In this paper, we also assume that there are always data to be sent, either because new data are being acquired or because there is a buffer which stores data until the data can be sent. In the latter case, a feedback channel is involved.

The goal in this paper is to maximize the total expected data returned during a mission or portion of a mission. This is the time integral of the total data rate chosen for transmission times the probability that reception could be carried out. For, we assume Shannon link coding, in which either the signal-to-noise ratio can sustain the particular code chosen, or the word and bit error probability are substantial and so the decoded data must be discarded.

Here we have a risk-balancing problem. If we choose too *high* a data rate, we risk losing all, but if we choose too *low*, we risk not getting as much out of the mission as we could have. We will consider the possibility of using simultaneous data rates, in an optimum joint coding scheme from the theory of broadcast channels. We shall also point out that mixed strategies, ones in which we randomly choose data rates, should not be used if the criterion is maximizing the expected data return, as it is here.

We may also be required to maintain a certain *minimum* data rate with high probability such as 0.99, for example, spacecraft health data and minimal imaging coverage. We call this the *base* data rate. The additional data that we may be able to receive is called *bonus* data.

Here there are three possibilities. The weather may turn out good enough to support the bonus data rate and the base data rate. Or the weather may not be that good, but may still support the base data rate. Finally, the weather may be so bad that we receive nothing. We still wish to maximize the expected total data returned. But we have the constraint that the probability that we receive nothing (this is the probability that the base data cannot be received) be held to a given small probability. This approach will be contrasted with the more traditional approach of using a single data rate only, but putting in a "weather pad" so that the chosen single data rate can be received with acceptably high probability.

## II. No Minimum Rate Required

Here we define $p(x)$ as the probability density that the weather is good enough to support data rate $x$. If the spacecraft were close enough to Earth, $p(x)$ would be a delta function because we know what the weather is. For a distant spacecraft, $p(x)$ represents the residual uncertainty after our best weather prediction is made. The density $p(x)$ is assumed known.

We will first choose a single rate deterministically depending on $p$. Later we will show that multiple rate strategies should not be considered here. This means no mixed strategies and no strategies employing more than one rate at once. But, if we need to retain a certain minimum rate with a high probability, we shall see in the next section that a dual-rate system should be employed. Here we ask, at what data rate $x_0$ should we choose to transmit so as to maximize the expected data return?

We receive $x_0$ bits per second with probability $P(x_0)$ given by

$$P(x_0) = \int_{x=x_0}^{x_{max}} p(x)\,dx \qquad (1)$$

Here $x_{max}$ is the data rate supportable in clear dry weather. We receive $x_0$ bits/sec with probability $P(x_0)$, and 0 bits/sec with probability $1 - P(x_0)$. This is because if the weather we actually have will not support the data rate that was transmitted, then according to Shannon we get nothing.

What is the expected data rate $E(x_0)$? It is just

$$E(x_0) = x_0 P(x_0) \qquad (2)$$

according to what we just said. So we want to choose an $x_0$ to maximize (2). This is the best we can do. In reasonable cases, we can use differentiation to maximize (2). The condition for a maximum is

$$P(x_0) - x_0 p(x_0) = 0 \qquad (3)$$

For example, let $p(x)$ be given by

$$p(x) = \frac{3}{2x_{max}^3}\left(x_{max}^2 - (x_{max} - x)^2\right) \text{ for } 0 \leqslant x \leqslant x_{max},$$

so that

$$P(x_0) = \frac{3}{2}\left(1 - \frac{x_0}{x_{max}}\right) - \frac{1}{2}\left(1 - \frac{x_0}{x_{max}}\right)^3.$$

This is the quadratic density function having its peak at $x_{max}$ and having the value 0 at $x = 0$. Equation (3) becomes

$$\frac{3}{2}\left(1 - \frac{x_0}{x_{max}}\right) - \frac{1}{2}\left(1 - \frac{x_0}{x_{max}}\right)^3 = \frac{3}{2}\frac{x_0}{x_{max}}\left[1 - \left(1 - \frac{x}{x_{max}}\right)^2\right]$$

We let $u = 1 - (x_0/x_{max})$ and solve this equation numerically to find $u \doteq 0.459$. Or,

$$\frac{x_0}{x_{max}} \doteq 0.541$$

The maximum expected data return is then

$$E(x_0) = x_0 P(x_0) = 0.346\, x_{max}$$

Suppose now that we were prescient and knew the actual weather $x$ in advance. We would transmit exactly at the rate $x$. The expected data return $\overline{E}$ with prescience is then

$$\overline{E} = \int_{x=0}^{x_{max}} x p(x)\,dx \qquad (4)$$

This of course must be greater than the value of $E(x_0)$ for the maximizing $x_0$. But, how much greater?

In the example, we can find $\bar{E}$. The result is

$$\bar{E} = (5/8)\, x_{max}$$

The gain in expected data return will be

$$\frac{\bar{E}}{E(x_0)} = \frac{(5/8)\, x_{max}}{0.346\, x_{max}} = 1.806 = 2.6 \text{ dB}$$

This is equivalent to a 2.6-dB power increase, at least if we have an infinite-bandwidth gaussian channel at our disposal. We lose 2.6 dB by not being able to predict the weather perfectly in this example. We lose in two ways. Some of the time we could have gotten rates higher than $x_0$ when we only got $x_0$, and some of the time we could have gotten something (less than $x_0$), instead of getting nothing.

How large can the loss be for not being able to predict the future perfectly? Appendix A shows that we can lose arbitrarily much. This is not surprising. Appendix A also shows that if the density $p(x)$ is nondecreasing in the interval of rates $[0, x_{max}]$, as is sometimes reasonable, then there is a maximum loss. It is 3 dB, attained only for the uniform density on $[0, x_{max}]$. For every other nondecreasing density, the loss will be less than 3 dB. The 2.6-dB loss in the quadratic density example is near the maximum possible for a nondecreasing density function.

Let us close this section by turning to another matter. We want to show that we do not need to consider multiple rates, either simultaneously or as part of a single-rate mixed strategy involving several rates.

It is easy to rule out mixed strategies. For, if we choose rates $x_i$ with probability $p_i$, the expected data return is $E(\{p_i\})$ given by

$$E(\{p_i\}) = \sum_i p_i \cdot x_i P(x_i) \tag{5}$$

We may as well always choose an $x_i$, say $x_i'$, maximizing $x_i \cdot P(x_i)$ over all $i$. This will yield a larger $E(\{p_i\})$ (unless all $x_i \cdot P(x_i)$ with $P_i \neq 0$ are equal). Mixed strategies do not pay.

Why should we not send data simultaneously at more than one rate if there is no minimum data rate required? We might think to do this to salvage something even if the weather turns out worse than we expected. Suppose we send data simultaneously at rates $x_i$, $1 \leqslant i \leqslant n$. We will use as our simultaneous coding scheme the degraded gaussian broadcast channel model of Ref. 1. Here increasing index means worse weather.

We see that all rates $x_j$ with $j \geqslant i$ can be received if weather $i$ holds at time of receipt. We shall let $u_i \geqslant x_i$ be the single-channel rate that could have been supported with weather $i$. Here the data $x_i$ are disjoint, so that we really do get credited for

$$\sum_{j>i} x_j$$

when weather $i$ occurs. If we want to maximize expected data return, there is no reason to repeat data on separate channels.

Actually, the total data returned if weather $i$ occurs cannot exceed the data that would have been returned had we put all our power into a single channel:

$$\sum_{j \geqslant i} x_j \leqslant u_i \tag{6}$$

This is because $u_i$ is the channel capacity when weather $i$ occurs. Reference 1 shows that the inequality in (6) is strict unless the bandwidth is infinite.

The probability that the weather is at least as good as $i$ is given by $P(u_i)$ from Eq. (1). The expected data returned when we use the multiple rates is then given by say $E_1(x_1, x_2, \ldots, x_n)$, where

$$E_1(x_1, x_2, \ldots, x_n) = \sum_{i=1}^{n} x_i P(u_i) \tag{7}$$

This is because the probability that $x_i$ can be received is the probability $P(u_i)$ that the weather is at least as good as $i$.

Because increasing $i$ means worse weather,

$$P(u_i) \leqslant P(u_2) \leqslant \ldots \leqslant P(u_n)$$

So, Eq. (7) can be converted into an inequality:

$$E_1(x_1, x_2, \ldots, x_n) \leqslant P(u_n) \sum_{i=1}^{n} x_i \tag{8}$$

In view of (6) with $i = 1$, (8) becomes

$$E_1(x_1, x_2, \ldots, x_n) \leqslant u_n P(u_n) = E(u_n) \tag{9}$$

The right-hand side of (9) is, by (2), the expected data return if we use the single rate $u_n$. This means we assume the worst weather and use all our power with the single rate $u_n$ tailored for that weather. This shows that single rates are best when there is no minimum data rate to be maintained.

## III. Minimum Rate Required

Suppose there is a minimum or base data rate $x_2$ that we strive to receive even under adverse conditions. This means that for some small $\epsilon > 0$, we want the probability of not being able to receive at least this base data to be $\epsilon$ or less. Typically, $\epsilon$ may be 0.05 or 0.01, but may even be 0. This is because there generally is a worst weather loss. For example, at X-band (8.5 GHz), the loss due to attenuation and the resulting noise temperature increase is at worst 10 dB or so into a roughly 30 K receiving system. We should and do assume that $\epsilon$ and $x_2$ are such that, in weather corresponding to probability $\epsilon$, if we put all our power into a single rate, a rate $x_2$ can in fact be supported. We call $x_2$ the *base* data rate, and the data to be sent at that rate is called *base* data.

As in Section II, we need only consider the following strategies. We choose a single other level, say $\eta$, $1 \geqslant \eta \geqslant \epsilon$, of weather degradation (not precluding that it be the same as $\epsilon$-weather). We transmit at the maximum rate $x_1$ such that in $\eta$-weather, the simultaneous rates $(x_1, x_2)$ can be supported. We call $x_1$ the *bonus* data rate, and the data communicated by it *bonus* data.

Here the problem is to choose $x_1$ (or alternatively $\eta$, for $x_1$ and $\eta$ determine each other when $x_2$ and $\epsilon$ are given) so as to maximize the total expected data return, base and bonus. There are three possibilities. The weather may be worse than $\epsilon$-weather, and we receive nothing. The weather may be at least as good as $\epsilon$-weather, but worse than $\eta$-weather. We then receive only base data. Finally, the weather may be at least as good as $\eta$-weather. We then receive base and bonus data.

Let us use broadcast channel theory from, for example, Ref. 1. We let $B$ be the bandwidth available for signalling. $B$ may be infinite, but the notation assumes that $B$ is finite. Also let $P$ be the total available received signal power. Let $N_1$ be the receiver noise power spectral density corresponding to $\eta$-weather, the weather that determines $x_1$. Let $N_2 \geqslant N_1$ be the noise density corresponding to $\epsilon$-weather. Then from Ref. 1 the allowable simultaneous rates $(x_1, x_2)$ at which we can reliably communicate are given by

$$\begin{aligned} x_1 &= B \log_2 \left(1 + \frac{\alpha P}{N_1 B}\right) \\ x_2 &= B \log_2 \left(1 + \frac{\bar{\alpha} P}{\alpha P + N_2 B}\right) \end{aligned} \tag{10}$$

Here $\alpha$ is a parameter between 0 and 1, and $\bar{\alpha} = 1 - \alpha$. Because $N_2$ is given (it is determined by $\epsilon$), and $x_2$ is a given requirement, $\alpha$ is determined from the second equation of (10). We note that as $B \to \infty$, (10) just becomes the time-division or frequency-division multiplexing formula. We shall not explore this here. Reference 2 considered this case when the weather distribution consisted of a finite number of delta functions.

The meaning of (10) is that we devote a fraction $\alpha$ of the power to the bonus data at rate $x_1$, ignoring the base data. We code for the resulting gaussian channel of signal power $\alpha P$, bandwidth $B$, and noise density $N_1$. The codewords for the base data at rate $x_2$ are assigned the remaining average power, which is $\bar{\alpha} P$. These "base" codewords are centered as a "cloud" around the "bonus" codewords. The bonus codewords look like Gaussian noise of power $\alpha P$ to a receiver designed for $\epsilon$-weather. We cannot decode them, and they add $\alpha P$ to the $N_2 B$ noise power denominator of the second equation of (10). But, if the weather is at least as good as $\eta$-weather, we *can* reliably decode the bonus codewords, remove them as noise by subtraction, and certainly then decode the base codewords. In this way, we see that we can do at least as well as (10), and Ref. 1 shows we can do no better. Our assumption that we can at least get rate $x_2$ in $\epsilon$-weather if we let $x_1$ be 0 means that these equations are not contradictory – no $\log_2$'s of negative numbers.

What is $\alpha$? The second equation of (10) becomes

$$2^{x_2/B} = 1 + \frac{\bar{\alpha} P}{\alpha P + N_2 B}$$

from which we may derive

$$\begin{aligned} \bar{\alpha} &= \left(1 + \frac{N_2 B}{P}\right) \left(1 - 2^{-x_2/B}\right) \\ \alpha &= 1 - \left(1 + \frac{N_2 B}{P}\right) \left(1 - 2^{-x_2/B}\right) \end{aligned} \tag{11}$$

and,

$$\alpha P = P - (P + N_2 B)(1 - 2^{-x_2/B}) \qquad (12)$$

Finally, the supportable bonus rate $x_1$ is given in terms of $N_1$ as

$$x_1 = B \log_2 \left(1 + \frac{P - (P + N_2 B)(1 - 2^{-x_2/B})}{N_1 B}\right) \qquad (13)$$

Equation (13) allows us to translate the probability distribution of the weather, which we may think of as a distribution on $N_1$, into a probability distribution on the attainable bonus data rate $x_1$.

We now observe that the only rate varying is $x_1$, because $x_2$ is fixed. So as far as optimization problems go, this is really the same problem as in the preceding section. We have a probability density $p(x_1)$ on rates $x_1$, and want to maximize the total expected bonus data returned. For, by our assumptions, the expected base data returned will be $(1 - \epsilon) x_2$, no matter what $x_1$ we choose. (This really assumes some continuity in the density function $p$, but we shall say no more about this.)

Let us now determine the maximum penalty due to lack of prescience, as we did in the previous section when no base data was required. We shall again want to assume that the probability density $p(x_1)$ on the achievable bonus data rate $x_1$ is nondecreasing in $x_1$. However, since we really would be given a probability density $r(N_1)$ on the noise density $N_1$ that actually occurs, this condition may be hard to check. Appendix B shows that if $N^2 r(N)$ is nonincreasing in its interval of definition $[N_{min}, N_{max}]$, then the density $p(x_1)$ will indeed be nondecreasing in its interval of definition $[0, x_{max}]$.

Under this mild condition, the previous section shows that the *rate* penalty in $x_1$ for not being prescient is at most 3 dB. What is this as a *power* penalty? This will help tell us how hard we should try to reduce the weather uncertainty.

We are really asking, by how much does the received power $P$ have to increase the double $x_1$? We still must have base rate $x_2$ achieved with probability $1 - \epsilon$. So, Eq. (10) holds, and (13) holds, with $2x_1$ replacing $x_1$ and say $P'$ replacing $P$. Starting at various forms of (13), we quickly realize that under the proper circumstances, doubling $x_1$ can mean an arbitrarily large increase ratio for $P$, or an arbitrarily many dB power penalty for not being prescient.

But, let us assume, as is reasonable, that $x_1/B$ in bits/cycle is not too large before the prescience doubling. This is probably a reasonable implementation constraint for most coding schemes that would be adopted. Let us agree that $x_1$ was equal to $B$ before the doubling. What happens to $P$ now?

Equation (13) can be written

$$2^{x_1/B} - 1 = \frac{P}{N_1 B} \cdot 2^{-x_2/B} - \frac{N_2}{N_1}(1 - 2^{-x_2/B}) \qquad (14)$$

Here, of course, $N_2 > N_1$. If $x_1/B = 1$, so that $2^{x_1/B} - 1 = 1$, we will have $2^{2x_1/B} - 1 = 2^2 - 1 = 3$. From (14), we need to at most *triple* $P$ to double $x_1$, when $x_1/B$ originally equalled 1. There is never more than a 4.8 dB power penalty for lack of prescience when we have a base data rate $x_2$ to maintain with probability $1 - \epsilon$, if the bonus rate $x_1$ does not exceed the available bandwidth $B$.

However, there is a more important question we can ask. What do we gain by the optimal two-rate base and bonus strategy over the more traditional approach of using a single rate? The single rate we should compare to is the one we get if we assume noise density corresponding to $\epsilon$-weather, the noise density $N_2$. For, this is the only way we can guarantee that we receive the base data in a single-rate strategy. The next section considers the gain of the optimal two-rate strategy over the traditional one-rate strategy.

## IV. Tradition Does Not Pay

The traditional single-rate strategy provides a data rate $x_s$ at least equal to $x_2$, because we put all our power into a single channel. Because the noise density $N_2$ is the one for $\epsilon$-weather, $x_s$ will be given by

$$x_s = B \log_2 \left(1 + \frac{P}{N_2 B}\right) \qquad (15)$$

We obtain $x_s$ with probability $1 - \epsilon$, and no data at all with probability $\epsilon$.

The expected total data returned by the traditional strategy will be $\overline{E}_s$, which is given by

$$\overline{E}_s = (1 - \epsilon) x_s \qquad (16)$$

With the optimal dual-rate strategy, the total data return $\overline{E}_{opt}$ is given by

$$\overline{E}_{opt} = \max_{x_1} x_1 P(x_1) + (1 - \epsilon) x_2 \tag{17}$$

Here $P(x_1)$ is the probability that the weather is good enough to support the bonus rate $x_1$ while the second channel is providing the base rate $x_2$.

We may ask, how large can $S = \overline{E}_{opt}/\overline{E}_s$ be? This of course depends on the weather density $r(N_1)$. But, it will always exceed 1 by definition of "optimum."

It is easy to show that $\overline{E}_{opt}/\overline{E}_s$ can be as large as we please if we pick the right parameters. For this, we should have a substantial probability, say around 1/2, that the noise $N_1$ is much smaller than $N_2$. But what if there is a minimum noise $N_{1,min}$? What is the largest that the data rate gain $S = \overline{E}_{opt}/\overline{E}_s$ can be in this case?

We want to choose a probability distribution on the noise to maximize the data rate gain $S$. From (16) and (17), we want to maximize

$$S = \frac{1}{(1 - \epsilon) x_s} \max_{x_1} x_1 P(x_1) + \frac{x_2}{x_s} \tag{18}$$

by choosing the distribution of receiver noise. Here $\epsilon$ and $x_2$ are given and

$$pr(N_1 > N_2) = \epsilon,$$

or

$$\int_{N_2^+}^{\infty} r(N)\, dN = \epsilon \tag{19}$$

Finally, $x_s$ is given according to Shannon by (15).

From (13), we find $x_1$ in terms of $N_1$ and $N_2$. The probability that the bonus data can be received is given by

$$P(x_1) = \int_0^{N_1} r(N) dN$$

Since $x_2$ is also given, the only term varying in (18) is $\max_{x_1} x_1 P(x_1)$.

How can we maximize this maximum on $x_1$ by the choice of the probability distribution on the noise density? Here $x_1$ is considered fixed. So we should maximize $P(x_1)$. Since the probability that base data cannot be received is $\epsilon$, $P(x_1)$ can be arbitrarily close to $1 - \epsilon$. We have to say "arbitrarily close" because we need to keep some probability just to the left of $N_2$ to "lock in" the base rate $x_2$.

The *max* of $x_1 P(x_1)$, which is really a sup, is then $(1 - \epsilon) x_1$. By arranging for a near-delta function of probability almost $1 - \epsilon$ at (just to the left of) $N_1$, we can indeed arrange that this $x_1$ be the optimum for the resulting distribution. This distribution, while somewhat artificial, is not too out of line, for the real weather distribution may tend to have a delta function of reasonable probability around the minimum noise, corresponding to clear dry weather. In this case, it also corresponds to $N_1$, because here $N_1 = N_{1,min}$.

We then find the maximum gain ratio $S$ given $N_1$ and $N_2$ from (18) as

$$S_{max}(N_1, N_2) = \frac{x_1 + x_2}{x_s} \tag{20}$$

Given the base rate $x_2$, the received power $P$, and the bandwidth $B$, we may attempt to maximize (20) over all possible $N_1$ and $N_2$ with $N_1 < N_2$, using (13) for $x_1$ and (15) for $x_s$. This is not very instructive.

However, for infinite bandwidth we can find the maximum in (20). We could do this easily from scratch without going through the finite-bandwidth case, but we shall take the limit as $B \to \infty$ in the finite-bandwidth case. We define a parameter $\rho$ greater than 1 as

$$\rho = \frac{N_2}{N_1} \tag{21}$$

Here $\rho$ is the ratio of the noise density in $\epsilon$-weather to that in clear dry weather, about 10 dB for low-noise X-band reception. Also let the parameter $\gamma$ less than 1 be defined as

$$\gamma = \frac{x_2}{x_s} \tag{22}$$

This is the ratio of the base data rate $x_2$ to the data rate that could be supported in $\epsilon$-weather if we put all our power into a single rate $x_s$.

From (13), for infinite bandwidth,

$$x_1 = \frac{P}{N_1 \ln 2} - x_2 \frac{N_2}{N_1} \tag{23}$$

From (15), we similarly have

$$x_s = \frac{P}{N_2 \ln 2} \tag{24}$$

Equation (20) then becomes

$$S_{max}(N_1, N_2) = \frac{\left(\frac{P}{N_1 \ln 2} - x_2\left(\frac{N_2}{N_1} - 1\right)\right)}{P/N_2 \ln 2}$$

$$S_{max}(N_1, N_2) = \rho - \gamma(\rho - 1) \tag{25}$$

We can check (25) by noting that

$$\text{as } \gamma \to 1, S_{max}(N_1, N_2) \to 1.$$

This means that there can be no savings in using the optimal strategy if the link can barely support the base rate in $\epsilon$-weather. This is as must be. But if, for example, $\gamma$ were 1/2, corresponding to $\epsilon$-weather being able to support *twice* the base rate,

$$S_{max}(N_1, N_2) = \rho - \frac{1}{2}(\rho - 1) = \frac{\rho + 1}{2}$$

If $\rho = 5$ for the given $\epsilon$, as may be typical of X-band for deep-space use, we can save up to a factor of 6/2 = 3 or 4.8 dB if the $\epsilon$-weather link could have supported twice the base rate. Here the probability that the noise is at least 5 times as bad as for clear dry weather is $\epsilon$. Half the power is devoted to base data, and half to bonus. The bonus rate is 5 times the base rate. There is five times as much energy devoted to each base bit as to each bonus bit.

In the next section, we do a more realistic case, one in which the bandwidth is finite and in which the probability distribution approximates the real ones that seem to occur for X-band weather.

## V. A Realistic Case

In this section, we adopt a weather model much like the actual weather statistics for deep-space communication at X-band. However, the results should be used only as guidelines and not for mission planning, because these are not actual weather statistics.

We consider a probability distribution varying from $N_{min} = 1$ as a normalization to $N_{max} = 10$, corresponding to the approximate 10 dB maximum X-band weather loss. We shall allow a $\delta$-function at $N = 1$, corresponding to a positive probability that the weather is perfectly clear and dry. Let the $\delta$-function have probability $1 - \beta$, so that the area of the continuous part $r_1(N)$ is $\beta$. We shall find $\beta$ to qualitatively match X-band weather statistics in this example.

If we agree that a 3-dB loss ($N = 2$) occurs 5% of the time, we can find $\beta$ under the assumption that the continuous part of the distribution is part of a parabola with its minimum at $N = 10$. Let the continuous part be $c(10 - N)^2$. Then

$$c \int_{N=1}^{10} (10 - N)^2 \, dN = \beta$$

and so

$$c = \frac{\beta}{243}$$

Then we have for $r_1(N)$

$$r_1(N) = \frac{\beta}{243}(10 - N)^2$$

The 3-dB loss occurs with probability 0.05. Or

$$\frac{\beta}{243} \int_{N=2}^{10} (10 - N)^2 \, dN = 0.05 = \frac{\beta}{729}(0.8)^3,$$

and

$$\beta = 0.0712, 1 - \beta = 0.9288$$

The overall weather density becomes

$$r(N) = 0.9288\, \delta(N - 1) + 2.930 \times 10^{-4} (10 - N)^2 \tag{26}$$

So the probability of clear weather is about 93%. This is not contrary to experience.

We have yet to choose an $\epsilon$. Rather than setting the criterion directly on $\epsilon$, we let it correspond to a 7-dB (factor of 5) loss, whereby $N_2 = 5$. Or

$$\epsilon = \int_5^{10} r_1(N)\, dN = 2.930 \times 10^{-4} \times \frac{5^3}{3} = 0.0122$$

This is slightly more than 1%. We want a 98.8% probability of getting at least the base data. One percent is lower than what is sometimes stated as a requirement in present deep space designs, but is probably close to the real requirement. The only reason 1% isn't required now is that the data rate penalty over clear weather would be too large (7 dB) with the traditional strategy. We shall see that we can pick up almost 5 dB of the 7 dB with the optimal dual-rate strategy.

We shall assume that the base-data-rate-to-bandwidth ratio $x_2/B$ is 1/8, and that, if we put all the power into a single channel at the $\epsilon$-weather noise density $N_2$ (=5), we could support a data rate of $2x_2 = x_s$ instead of the $x_2$ we will actually get. Or, from the channel capacity formula (15),

$$2x_2 = B \log_2 \left(1 + \frac{P}{N_2 B}\right)$$

$$2^{2x_2/B} - 1 = \frac{P}{N_2 B}, 2^{1/4} - 1 = \frac{P}{N_2 B} \tag{27}$$

So, $P/N_2B = 0.1892$, and, since $N_2 = 5$, $P/B = 0.9460$. We have determined the power-to-bandwidth ratio $P/B$. If we had kept better track of units, the dimensions would be joules/cycle (or watts/Hz).

From (11), we can now find

$$\overline{\alpha} = \left(1 + \frac{1}{0.1892}\right)(1 - 2^{-1/8}) = 0.522, \alpha = 0.478$$

We put 52.2% of the power into the base ($x_2$) channel, and 47.8% into the bonus ($x_1$) channel. But, we still need to find $x_1$ and $N_1$. From (10), with $P/B = 0.9460$, we have

$$x_1 = B \log_2 \left(1 + \frac{(0.478)(0.9460)}{N_1}\right)$$

$$\frac{x_1}{B} = \log_2 \left(1 + \frac{0.4522}{N_1}\right) \tag{28}$$

As an aside, we remark that, within $\overline{\alpha} = 0.522$, the power to the base channel is reduced by 0.522. And, from (10), the noise density is increased, due to bonus coding, by

$$\left(\frac{\alpha P + N_2 B}{B}\right) - N_2 = \alpha P/B = 0.4522,$$

the same 0.4522 of Eq. (28). This is a percentage increase of $0.4522/N_2 = 0.4522/5 = 0.905^- = 9.0\%$. The noise density in the base channel increases 9.0% due to the bonus codewords being seen as noise by the base decoder.

The noise density increase fraction is, from (10), (11), and (27), actually

$$\frac{\alpha P}{N_2 B} = \left(1 + \frac{P}{N_2 B}\right) 2^{-x_2/B} - 1 = 2^{x_2/B} - 1$$

It depends only on the assumed $x_2/B$ and $x_s/B$, and not on the probability density $r(N)$ chosen nor on $\epsilon$. The power devoted to the base channel drops by a factor of 0.522 due to bonus coding, and the noise increases by a factor of 1.0905. So, the signal-to-noise drops by a factor of $0.522/1.09 = 0.479^-$. This is, of course, the drop in signal-to-noise ratio which exactly cuts a data-rate-to-bandwidth ratio of $2x_2/B = 1/4$ to one of $x_2/B = 1/8$:

$$\frac{2^{1/8} - 1}{2^{1/4} - 1} = 0.478^+$$

The probability that the bonus rate can be supported if we design it for weather producing noise density $N_1$ is say $P_0(N_1)$, where

$$P_0(N_1) = 1 - \int_{N_1}^{10} r_1(N)\, dN = 1 - \frac{(10 - N_1)^3}{729}$$

We seek to maximize, by choice of $N_1$, the function $x_1 P_0(N_1)$, or, by (28), maximize

$$G(N_1) = (\ln 2)\, F(N_1) = \left[\ln \left(1 + \frac{0.4522}{N_1}\right)\right] \left(1 - \frac{(10 - N_1)^3}{10240}\right)$$

The derivative $G'(N_1)$ is

$$G'(N_1) = \frac{1}{\left(1 + \frac{0.4522}{N_1}\right)} \left(\frac{-0.4522}{N_1^2}\right) \left(1 - \frac{(10 - N_1)^3}{10240}\right)$$

$$+ \left[\ln \left(1 + \frac{0.4522}{N_1}\right)\right] (-3) \frac{(10 - N_1)^2}{10240}$$

Both terms are clearly negative in $1 \leqslant N_1 < 10$, so the maximum is at $N_1 = 1$. We play for clear weather. Because the weather is clear with probability 0.9288, this is no surprise.

From (28), we see that

$$\frac{x_1}{B} = \log_2 (1 + 0.4522) = 0.5388,$$

$$\frac{x_1}{B} P_0(N_1) = (0.5388)(0.9288) = 0.5004$$

From (17), then,

$$\frac{\overline{E}_{opt}}{B} = 0.5004 + (1 - \epsilon)(1/8) = 0.5004 + \frac{0.9878}{8},$$

$$\frac{\overline{E}_{opt}}{B} = 0.6239.$$

Compared with the $2x_2/B$ of 1/4, we have a gain of a factor of 2.50 or 4.0 dB in total expected data rate, and a factor of 2.86 or 4.6 dB in power.

We note that the adopted optimum strategy gives a data-rate-to-bandwidth total of $0.5388 + 0.1250 = 0.6638$ with probability $1 - \beta = 0.9288$, a data-rate-to-bandwidth total of 0.1250 with probability $0.9878 - 0.9288 = 0.0590$, and no data with probability 0.0122. The traditional strategy would give 0.2500 with probability 0.9878 and no data with probability 0.0122. We only get half as much data in the $\epsilon = 0.0122$ weather, but the enormous gain in $N_1 = 1$ (clear dry) weather much more than makes up for this on the average. The optimal dual-rate strategy gains a 2.5 factor increase in expected data rate, which is equivalent, when the finite bandwidth is taken into account, to a power gain of 4.6 dB. Tradition does not pay.

## References


1. Cover, Thomas M., and Abbas El Gamal, "Multiple User Information Theory," *Proc. IEEE*, Vol. 68, 1980, pp. 1466-1483, esp. Sec. II C.
2. Butman, Stanley "Transmission Strategies for Atmospheric Entry Probes," *Proc. 1972 National Telemetering Conf.* (NTC 72), IEEE Publication, pp. 28E-1 to 28E-4.

# Appendix A
## Penalty for Lack of Prescience

Here we answer the question raised in Section II. What is the maximum loss in expected data return from not being able to know what the weather will be?

If there is no restriction on $p(x)$, the loss can be arbitrarily large. For example, let only integer rates $j = 2, 3, \ldots$ be supportable. The example can be modified so that the probability distribution of rates has a continuous density and a maximum rate. The conclusion will be the same.

Suppose the probability $p_j$ that the actual weather would support rate $j$ is

$$p_j = \frac{c}{j^2}, \qquad j = 2, 3, \ldots$$

Here $c$ is an irrelevant constant (actually $1/[(\pi^2/6) - 1]$. For this weather distribution

$$\overline{E} = \sum_{j=1}^{\infty} jp_j = \sum_{j=2}^{\infty} \frac{c}{j} = \infty$$

But the expected data return $E_i$ if we chose rate $i$ is

$$i \sum_{j=i}^{\infty} p_j = ci \sum_{j=i}^{\infty} \frac{1}{j^2} < ci \int_{x=i}^{\infty} \frac{1}{(x-1)^2}\, dx$$

$$E_i < \frac{ci}{i-1} \leqslant 2c$$

So the $i$ maximizing $E_i$, which we may call $i_0$, has

$$E(i_0) < 2c$$

But the expected data return $\overline{E}$ with prescience is $\overline{E} = \infty$. There is an infinite penalty for lack of prescience. This is as expected.

Now we ask, suppose $p(x)$ must be nondecreasing on the interval $[0, x_{max}]$. What is the largest penalty for lack of prescience now? We may scale rates so that $x_{max} = 1$. If $p(x) = 1$ on $[0, 1]$, then

$$\overline{E} = \int_0^1 x\, dx = 1/2$$

Also,

$$x_0 P(x_0) = x_0(1 - x_0)$$

has maximum $1/4$ at $x_0 = 1/2$. For this $x_0$, $E(1/2) = 1/4$ and the penalty is

$$\overline{E}/E(1/2) = \left(\frac{1/2}{1/4}\right) = 2 = 3 \text{ dB}$$

We shall show that the 3-dB loss encountered above is the worse case when $p(x)$ is nondecreasing. In fact, we shall show that if $\overline{X}$ is the mean or expected data rate given by

$$\overline{X} = \int_0^1 xp(x)\, dx$$

(which we have called $\overline{E}$ above) then using $\overline{X}$ as the actually chosen rate $x_0$ never loses more than 3 dB, and loses less unless $p(x)$ is the uniform density on $[x_{min}, x_{max}]$. We note that if $x_{min} > 0$, $\overline{X}$ will not be the optimum data rate to choose.

We seek to show

$$\frac{\overline{X}}{\left(\overline{X} \int_{\overline{X}}^1 p(x)\, dx\right)} \leqslant 2$$

when $p(x)$ is nondecreasing on $[0, 1]$, with equality only when $p(x)$ is constant on its interval of support. We want to show

$$\int_{\overline{X}}^1 p(x)\, dx \geqslant \frac{1}{2} \tag{A-1}$$

for $p(x)$ nondecreasing. Another way to put this is that the median $X_{1/2}$, the point such that half the probability lies to the right and half to the left, satisfies

$$\overline{X} \leqslant X_{1/2} \tag{A-2}$$

if $p(x)$ is nondecreasing, with equality only when $p(x)$ is constant for $x_{min} \leqslant x \leqslant x_{max}$.

Let $p(X_{1/2}) = h$. We *increase* $\overline{X}$ for the same $X_{1/2}$ if we replace the original $p(x)$ by one which is equal to the constant

$h$ for $b \leqslant x \leqslant X_{1/2}$ and equals 0 for $0 \leqslant x \leqslant b$. Here the constant $b$ is given by

$$h(X_{1/2} - b) = 1/2$$

$$b = X_{1/2} - \frac{1}{2h}$$

Note that $b \geqslant 0$. For the *maximum* probability to the left of $X_{1/2}$ is the area of the rectangle with base $X_{1/2}$ and height $h$. This probability is exactly 1/2, so

$$h \cdot X_{1/2} \geqslant \frac{1}{2}$$

$$b = X_{1/2} - \frac{1}{2h} \geqslant 0$$

What happens to the right of $\overline{X}$? We can increase $\overline{X}$ keeping $X_{1/2}$ the same by extending $p(x)$ to the right at the constant value $h$ for a length long enough, to $a$, say, with $a > 1$, to make the resulting rectangle have area 1/2 to the right of $X_{1/2}$. Figure A-1 explains these operations.

The resulting uniform density on $[b, a]$ has a *larger* mean than the original $p(x)$, for probability has been pushed to the right. However, it has the *same* median $X_{1/2}$. But the resulting uniform distribution has mean $\overline{X}'$ equal to $X_{1/2}$, for the mean of a uniform distribution is also its median. The result is the following:

$$\overline{X} \leqslant \overline{X}' = X_{1/2} \qquad \text{(A-3)}$$

This provides (A-2). Note that if the original $p(x)$ were not uniform on its interval of support, we would have actually increased $\overline{X}$ to $\overline{X}'$, so that $\overline{X} < \overline{X}'$. This shows that the 3-dB loss holds *only* for uniform $p(x)$. So 3 dB is the worst loss possible for a nondecreasing $p(x)$. This loss is attained only for uniform $p(x)$, and we can guarantee that we don't lose more than 3 dB from the prescience value by choosing to transmit at the rate $\overline{X}$, the mean data rate we could get with prescience. Of course, if we choose the $x_0$ maximizing (2), we will be able to cut the prescience loss still further in all cases except where $x_{min} = 0$ and $p(x)$ is uniform. But for the uniform distribution when the prescience loss is 3 dB, $x_0 = x_{max}/2$ is the best $x_0$.

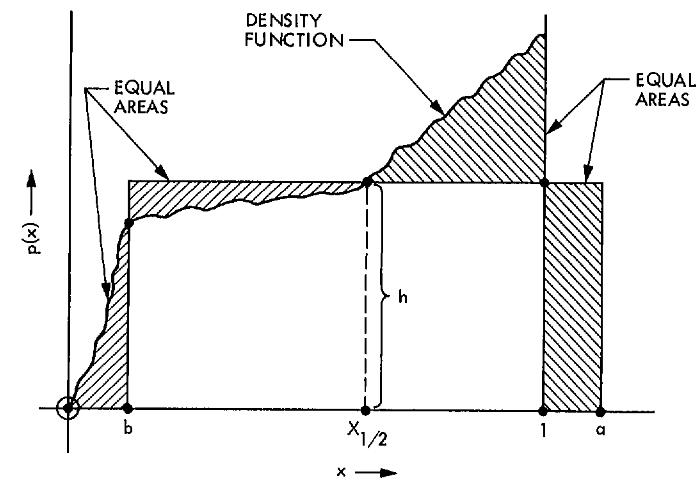


**Fig. A-1. Flattening *p*(*x*)**

# Appendix B
# Condition for Nondecreasing $p(x_1)$

Here we find a condition on the probability density $r(N_1)$ of the noise $N_1$ that implies that the probability density $p(x_1)$ of the bonus data rate $x_1$ is nondecreasing up to the maximum rate $x_{max}$. We referred to this condition in Section III.

Equation (13) can be written

$$x_1 = B \log_2 \left(1 + \frac{L}{N_1 B}\right) \tag{B-1}$$

Here

$$L = P - (P + N_2 B)(1 - 2^{-x_2/B})$$

is a positive constant, positive because the rate $x_1$ is a positive rate.

We may write

$$pr(X_1 \leqslant x_1) = pr(n_1 \geqslant N_1) = R(N_1) \tag{B-2}$$

Here $X_1$ is the random variable of rates $x_1$, $n_1$ is the random variable of corresponding noise densities $N_1$, and $R(N_1)$ is 1 minus the cumulative distribution of the noise random variable $n_1$ evaluated at the particular $\epsilon$-weather noise $N_2$.

We can differentiate (B-2) with respect to $x_1$, and use (B-1) to differentiate $N_1$ with respect to $x_1$. The result is

$$p(x_1) = \frac{d}{dx_1} R(N_1) = \frac{dN_1}{dx_1} \frac{d}{dN_1} R(N_1) = -\frac{dN_1}{dx_1} r(N_1) \tag{B-3}$$

But, from (B-1),

$$\frac{dN_1}{dx_1} = \frac{1}{dx_1/dN_1} = -\frac{\ln 2}{L}\left(N_1^2 + \frac{L}{B} N_1\right) \tag{B-4}$$

(B-3) becomes

$$p(x_1) = \frac{\ln 2}{L}\left(N_1^2 + \frac{L}{B} N_1\right) r(N_1) \tag{B-5}$$

If $N_1^2\, r(N_1)$ is nonincreasing in $N_1$, then $N_1\, r(N_1)$ is all the more nonincreasing in $N_1$. Since $L/B$ is nonnegative, the sum $N_1^2\, r(N_1) + (L/B)N_1\, r(N_1)$ is nonincreasing in $N_1$. Since increasing noise $N_1$ corresponds to a decreasing bonus rate $x_1$, (B-5) shows that $p(x_1)$ is nondecreasing in $x_1$ if $N_1^2\, r(N_1)$ is nonincreasing in $N_1$. (In fact, the condition on $p(x_1)$ is almost but not quite *equivalent* to the condition on $r(N_1)$.)